patient-specific or subpopulation-specific, such as ethnic, age, or gender groups, design or selection of drugs.

Please replace the paragraph on page 7, line 8-30, with the following:

A computer-based method for identifying compensatory mutations in a target protein is also provided. The method involves obtaining the amino acid sequence of a target protein containing multiple amino acid mutations that is expressed in a patient, where the structure of a form of the target protein that responds to a particular drug, including the active site, has been structurally characterized; generating a 3-D structural model of the mutated protein; comparing the structure of the mutated protein with the form of the protein that responds to the drug to identify structural differences and/or similarities arising from the mutations; comparing the biological activities of the drug against the mutated protein and the form of the protein that responds to the drug to determine the effects of the mutations on drug response; and identifying the mutations in the protein that affect biological activity based on the comparisons. The target biomolecules can also be used in a method referred to herein as computational phenotyping to predict drug sensitivity or resistance for a given genotype. These computer-based method for identifying phenotypes *in silico* are provided. The methods involve obtaining from a patient/specimen, such as a body fluid or tissue sample, including blood, cerebral spinal fluid, urine, saliva, sweat and tissue samples, the amino acid sequence of a target protein; generating a 3-D structural model of the target protein; performing protein-drug binding analyses; and predicting drug sensitivity or resistance based on the protein-drug binding analyses.

Please replace the paragraph on page 8, lines 24-31, with the following:

The databases can also be used for identification of invariant residues and regions of a target biomolecule, such as an HIV protease or reverse transcriptase. The identified invariant regions are then used to computationally screen compounds, preferably small molecules by assessing binding

interactions. The compounds so-identified serve as candidates for drugs that will be effective for a larger proportion of a population or against a broader range of variants of a pathogen, where the target protein is from a pathogen.

Please replace the paragraphs on page 12, lines 1-7, with the following:

As used herein, structural variant proteins refer the variety of 3-D molecular structures or models thereof that result from the polymorphisms. These variants typically arise from transcription and translation of genes containing genetic polymorphisms, but also include differentially glycosylated or otherwise post-translationally modified variants that potentially exhibit differential interactions with drugs and drug candidates.

Please replace the paragraph on page 12, lines 19-25, with the following:

As used herein, structure-based drug design refers to computer-based methods in which 3-D coordinates for molecular structures are used to identify potential drugs that can interact with a biological receptor. Examples of such methods include, but are not limited to, searching of small molecule libraries or databases, conformational searching of a ligand within an active site of identified biologically active conformations or computational docking methods.

Please replace the paragraph on page 13, lines 1-16, with the following:

As used herein, energetic refinement refers to the use of molecular mechanics simulation techniques, such as energy minimization or molecular dynamics, or other techniques, such as quantum-based approaches, to "adjust" the coordinates of a molecular structural model to bring it into a stable, low energy, conformation. In molecular mechanics simulations, the potential energy of a molecular system is represented as a function of its atomic coordinates along with a set of atomic parameters, called a force field. Energy minimization refers to a method wherein the coordinates of a molecular conformation are adjusted according to a target function that results in a lower energy conformation. Molecular dynamics refers to methods for simulating molecular motion by inputting kinetic energy into the molecular system corresponding to a

specified temperature, and integrating the classical equations of motion for the molecular system. During a molecular dynamics simulation, a system undergoes conformational changes so that different parts of its accessible phase space are explored.

Please replace the paragraphs on page 14, line 12, through page 15, line 18, with the following:

As used herein, haplotype refers to two or more polymorphism located on a single DNA strand. Hence, haplotyping refers to identification of two or more polymorphisms on a single DNA strand. Haplotypes can be indicative of a phenotype.

As used herein, a parameter is any input data that will serve as a basis for sorting the database. These parameters will include phenotypic traits, medical histories, family histories and any other such information elicited from a subject or observed about the subject. A parameter may describe the subject, some historical or current environmental or social influence experienced by the subject, or a condition or environmental influence on someone related to the subject. Parameters include, but are not limited to, any of those described herein, and known to those of skill in the art.

As used herein, computational phenotyping, refers to computer-based processes that assess the phenotype resulting from a particular genotype. The phenotype describes observables, such as, but are not limited to, the structure of the encoded protein, its functional morphological and structural attributes. In particular, as contemplated herein, the phenotype that is assessed is the interaction of a protein with a particular compound, particularly a drug. As exemplified herein, the method provides a means to select an effective drug for a particular subject, particularly mammals, or class thereof.

As used herein, a database refers to a collection of data; in this case data relating to polymorphic variants. Hence a database contains the nucleic acid sequences encoding the variants, or a portion of the variant, such as a portion containing the active site or targeted site. Additionally, the database may contain other information related to each entry, including but are not limited to, the corresponding 3-D structure of the encoded protein (or a portion thereof) and information regarding the source of each sequence. Some of the entries in a database may be identical, and for purposes herein, a database contains at least 2 different entries, typically far more than 2 entries. The number of entries depends upon the protein of interest and variety and number of polymorphisms that exist. Generally a database will have at least 10 different entries, typically more than 100, more than 500, more than 1000, more than 2000, 3000, 4000, 5000, 8000, 10,000, 50,000, 100,000 and greater. Databases herein containing 20,000 entries and more have been generated and are exemplified herein.

Please replace the paragraph on page 22, line 28, through page 23, line 9, with the following:

It is shown herein that it is advantageous to use 3-D molecular structures in drug design rather than to consider primary sequence alone. For example, most drugs target proteins either in the afflicted organism or in a pathogen. Disease, drug action and toxicity are all manifested at the protein level. Although the nucleotide sequences of genetic polymorphisms might appear to be quite different, the resulting protein targets may have similar shapes and, therefore, the protein's biological function might be the same. Conversely, although genetic polymorphism sequences might appear similar, the resulting proteins may have critical differences in their 3-D structures that greatly affect biological activity. Thus, use of 3-D protein structure models in such methods provide advantages not heretofore realized. Methods for generating 3-D structures are known to those of skill in the art and are also provided herein.

Please replace the paragraphs on page 24, lines 18-29, with the following:

The target gene is one that exhibits polymorphisms (i.e., sequence variations among a population) and the target protein is the product of a gene exhibiting genetic polymorphisms, or sequence variations, as described herein. Any gene or protein that exhibits polymorphisms is contemplated herein. In particular, genes that encode proteins, polypeptides, or oligopeptides that are targets for drug interaction are contemplated herein. The genetic polymorphisms can occur in the genes of pathogens (*e.g.* viruses, bacteriae, and fungi), parasites, plants, animals, and humans. As such, the sequence of a target protein can be obtained by the isolation and analysis of the gene or gene product in samples taken from pathogens, parasites, plants, animals, and humans, most preferably from humans.

Please replace the paragraphs on page 29, line 22, through page 30, line 30, with the following:

Once the conserved regions of the model are assembled, *ab initio* loop prediction (Dudek *et al.* (1998) *J. Comp. Chem. 19*:548-573) indicated at 106A or *ab initio* secondary structure generation techniques of block 106B, techniques in which the alignments are adjusted using information on the secondary structure, functional residues, and disulfide bonds as described herein, can be used to complete the model (e.g. U.S. Patents Nos. 5,331,573; 5,579,250; and 5,612,895). This model, complete with loops, is then subjected to refinement procedures (block 110) based on molecular mechanics, molecular dynamics, and simulated annealing methods. Energetic refinement of the structure can be accomplished by performing molecular mechanics calculations using, for example, an ECEPP type force field (Dudek *et al.* (1998) *J. Comp. Chem. 19*:548-573) or through molecular dynamics simulations using, for example, a modified AMBER type force field (Ramnarayan *et al.* (1990) *J. Chem. Phys. 92:7057-7076.* As known to those of skill in the art a modified AMBER (version

3.3) force field is a fully vectorized version of AMBER (3.0) with coordinate coupling, intra/inter decomposition, and the option to include the polarization energy as part of the total energy (see, *e.g.*, Weiner *et al.* (1986) *J. Comp. Chem. 7*:230-252). If necessary, the 3-D structures can be dynamically refined, for example, by using a simulated annealing protocol (*e.g.*,, 100 ps equilibration, 500 ps dynamics, up to 1000°K, 1 fs data collection).

The refinement process step 110 is used to offset problems that may arise when homology models are not built carefully or when they are built using fully automated methods. Problems that may arise include chain breaks (e.g. consecutive C^{α} atoms are farther apart than the optimum distance of 3.7 to 3.9 Å); distorted geometry (e.g. bond lengths and bond angles are too far from their optimal values); *cis*-peptide bonds (*e.g.*, incorrect isomerization of the peptide backbone in non-proline residues when it is not required); disallowed backbone and side-chain conformations (*e.g.*, dihedral angles do not satisfy the Ramachandran plot (see, Balasubramanian (1974) *Nature 266*:856-857) criteria for a fully favorable protein structure conformation); and misfolded loops (e.g. non-homologous loops are generated in unnatural conformations). The refinement procedure 110 removes distortions of covalent geometry by using energetic methods, converts disallowed backbone and side-chain conformations into allowed ones using simulated annealing methods, conserves protein core structure and secondary structural elements built by homology, and rebuilds unnatural loop constructions (Dudek *et al.* (1998) *J. Comp. Chem. 19*:548-573).

Please replace the paragraph on page 32, line 15, through page 33, line 2, with the following:

Next, at block 214, the 3-D structural models for all variants are generated. A refinement process is then completed at block 216 for the structural models. As noted above in connection with FIG. 1, the process involves subjecting each model, complete with loops, to refinement procedures

based on molecular mechanics, molecular dynamics, and simulated annealing methods. As before, the energetic refinement of the structure can be accomplished by performing molecular mechanics calculations using an ECEPP type force field (Dudek *et al.* (1998) *J. Comp. Chem. 19*:548-573), or through molecular dynamics simulations using, for example, a modified AMBER type force field (Ramnarayan *et al.* (1990) *J. Chem. Phys. 92*:7057-7076), where a modified AMBER (version 3.3) force field is a fully vectorized version of AMBER (3.0) with coordinate coupling, intra/inter decomposition, and the option to include the polarization energy as part of the total energy (Weiner et al. (1986), J. Comp. Chem. 7:230-252). If necessary, the 3-D structures can be dynamically refined, for example, by using a simulated annealing protocol (*e.g.*,, 100 ps equilibration, 500 ps dynamics, up to 1000°K, 1 fs data collection).

F11 cont'd

Please replace the paragraph on page 34, lines 1-9, with the following:

At block 228, once the models are determined to be satisfactory, drug molecules are docked with the structural variant models. Next, at block 330, the free energy of binding is evaluated with the potential drugs under study for each structural variant model. At block 332, the total free energy of binding is decomposed, based on the interacting residue in the protein active site. Lastly, at block 334, the free energy of binding is correlated with patient data, if the data is available. Thus, the 3-D structural data is employed in drug design. Details of using such structural data in drug design are described further below.

F12

Please replace the paragraph on page 34, lines 11-15, with the following:

The crystal structure of any protein can be determined empirically and the resulting coordinates used as the basis for determining structures of variants. Such structures are often known (see, *e.g.*, Kohlstaedt *et al.* (1992) *Science 256*:1773-1790 for a crystal structure of HIV-1 RT bound to a ligand).

F13

Pleas replac the paragraph on pag 38, line 13, through page 39, line 8, with the following:

New potential drug candidates can be designed by identifying potential small molecule drugs that can bind to a particular structural variant. This is accomplished, for example, by methods including, but are not limited to, methods for electronic screening of small molecule databases as described herein, methods involving modifying the functional groups of existing drugs *in silico*, methods of *de novo* ligand design. Methods for computationally designing drugs are known to those of skill in the art and include, but are not limited to, DOCK (Kuntz *et al.* (1982) "A Geometric Approach to Macromolecule-Ligand Interactions", J. Mol. Biol., 161:269-288; available from University of Ca, San Francisco); and AUTODOCK (see, Goodsell *et al.* (1990) "Automated Docking of Substrates to Proteins by Simulated Annealing", Proteins: Structure, Function, and Genetics, 8, pp. 195-202; available from Scripps Research Institute, La Jolla); GRID (Oxford University, Oxford, UK); CAVEAT (UC Berkeley, Ca), LEGEND (Molecular Simulations, Inc., San Diego, CA); LUDI (Molecular Simulations, Inc., San Diego, CA); HOOK (Molecular Simulations, Inc., San Diego, CA); CLIX (CSIRO, Australia); GROW (Upjohn Laboratories, Kalamazoo); others including HINT, LUDI, NEWLEAD, HOOK, PRO-LIGAND and CONCERTS (see, M. Murcko, "An Introduction to De Novo Ligand Design" in Practical Application of Computer-Aided Drug Design, Charifson, Ed., Marcel Dekker, NY, pp 305-354), methods based on QSAR (quantitative structure-activity relationships, *QSAR and Drug Design: New Developments and Applications*, Fugita, Ed., (1995) Elsevier, pp 3-81; 3D QSAR in Drug Design, Kubinyi, Ed., (1993) Escom, Leiden), and other methods known to those of skill in the art for determining molecules that have optimal binding interactions with a selected target.

Please replace the paragraph on page 39, line 15, through page 40, line 4, with the following:

After the computational docking step, the free energy of binding of the docked complex is calculated, and the total free energy of binding is decomposed based on the interacting residues in the protein active site or sites deemed important for protein activity. Analyses of the binding energies are needed to identify drug candidates. If needed or desired, the free energy of binding of different drugs or potential drugs to each structural variant model can be calculated by subtracting the free energy of the non-interacting protein and drug from the free energy of the protein-drug complex. The total free energy of binding is decomposed into its various thermodynamic components, e.g. enthalpic and entropic components, based on the interacting residues in the protein active site in a solvated model to characterize the structural and thermodynamic features in the mode of drug binding and to determine the contribution of the solvent (see, *e.g.*, Wang *et al.* (1996) *J. Am. Chem. Soc. 118*:995-1001; Wang *et al.* (1995) *J. Mol. Biol. 253*:473-492; Ortiz *et al.* (1995) *J. Med. Chem. 38*:2681-2691, which describes a computational method for deducing QSARs from ligand-macromolecule complexes). Following the computational drug design protocol described herein, any potential new drugs that are identified can be synthesized in, for example, industry or academia, and subjected to further biological testing, such as *in vitro* studies or pre-clinical and clinical *in vivo* testing.

Please replace the paragraph on page 44, lines 6-16, with the following:

If common structural features are observed over a range of protein targets that are derived from genetic polymorphisms, these common features may be used to design a drug that is effective with a variety of genetic polymorphisms and thus many patients. The retention of certain common structural features over a large number of genetic polymorphisms suggests that those features may not be mutable because the conserved structure may be essential to protein

function, *e.g.*, to the viability of an infectious organism or virus. Such conserved structural elements are prime targets for structure-based drug design, *e.g.*, anti-infective or antibiotic drug design, and can lead to highly effective therapies.

Please replace the paragraph on page 44, line 28, through page 45, line 14, with the following:

In comparing sets of related protein structures, such as those with the same biological function or those resulting from genetic polymorphisms, certain parts of the structural framework are often found to be conserved, while other parts vary among the proteins. Mutations that occur in the conserved regions of the structure can have significant effects biological activity. For example, in viruses, the conserved features can be essential to protein function and, thus, to the viability of the infectious organism or virus. Identifying the conserved structural features over a range of structures often gives insight into which structural features are necessary for biological activity and are therefore non-mutable. By analyzing a number of structural variants derived from genetic polymorphisms that exhibit drug resistance, it is possible to identify or design drugs that interact best with the common structural features in all of the variants. Using these features in structure-based drug design studies leads to the identification of drugs that retain biological activity despite multiple mutations, or polymorphisms, and could help to overcome the problem of drug resistance.

Please replace the paragraph on page 51, lines 10-29, with the following:

A database is preferably interfaced to a molecular graphics package that includes 3-D visualization and structural analysis tools, to analyze similarities and variations in the protein structural variant models (see, copending U.S. application Serial No. 09/531,995, which is published as International PCT application No. WO 00/57309, and is a continuation-in-part of U.S. application Serial No. 09/272,814, filed March 19, 1999). Briefly, International PCT

application No. WO 00/57309 provides a database and interface for access to 3-D molecular structures and associated properties, which can be used to facilitate the design of potential new therapeutics. The interface also provides access to other structure-based drug discovery tools and to other databases, such as databases of chemical structures, including fine chemical or combinatorial libraries, for use in structure-focused high-throughput screening, as well as to a host of public domain databases and bioinformatics sites. The interface also provides access to other structure-based drug discovery tools and to other databases, such as databases of chemical structures, including fine chemical or combinatorial libraries, for use in structure-focused high-throughput screening, as well as to a host of public domain databases and bioinformatics sites. This interface can be modified as needed to adapt for use with a particular database.

Please replace the paragraph on page 54, lines 20-29, with the following:

Databases containing data representative of the 3-D structure of structural variants encoded by a selected gene or genes or the 3-D structure of other polymorphic variants are provided. The selected genes can be genes of drug targets, such as receptors, and genes of infectious agents, such as the HIV protease or reverse transcriptase. Exemplary databases are presented in Example 5 which describes the construction, interface, use and applications of HIV PR and RT databases. These databases may be stored on any suitable medium and used in any suitable computer system. Systems and methods for generating, storing and processing databases are well known.

Please replace the paragraph on page 69, lines 20-25, with the following:

To modify these compounds, important pharmacophore features on the surface of the receptor that are critical for binding of the compounds were identified. These features include a hydrophobic belt, a hydrophilic belt and 3 hydrogen bond donor sites. A few potential hydrogen bonding sites, which are

F20 cont

not used by the current compounds, were also derived, and can be used for designing more potent binders.

Please replace the paragraph on page 76, line 30, through page 77, line 2, with the following:

F21

Computational or *in silico* phenotyping is performed to assess phenotypic properties of a protein. This example demonstrates application of this method to HIV-1 protease and reverse transcriptase to test whether the efficacy of various protease inhibitors for an HIV patient.

IN THE CLAIMS:

Please add claims 95-127 as follows:

F22

95. (New) The method of claim 1, further comprising:
after determining the 3-D structural variant models, exporting some or all of the models into a program that computationally docks the models with test compounds to assess intermolecular interactions.

96. (New) The method of claim 1, wherein the selected model structures represent structural variants derived from subjects who receive a specific treatment regimen.

97. (New) The method of claim 96, wherein a subject is a human.

98. (New) The method of claim 1, wherein the selected model structures represent structural variants derived from subjects who exhibit a particular clinical response to a given drug.

99. (New) The method of claim 98, wherein a subject is a human.

100. (New) The method of claim 1, wherein the selected model structures represent structural variants based on the duration of a particular drug treatment.

101. (New) The method of claim 1, wherein the step of determining 3-D protein structural variant models is performed by a method selected from the group consisting of experimental methods, searching protein structure databases, homology modeling, molecular modeling, de novo protein folding,

computational protein structure prediction, *ab initio* methods and combinations thereof.

102. (New) The method of claim 101, wherein the experimental methods include x-ray crystallography and NMR spectroscopy.

F22

103. (New) The method of claim 1, wherein the step of determining 3-D protein structural variant models is performed by a combination of homology modeling and *ab initio* methods.

104. (New) The method of claim 13, wherein the step of determining 3-D protein structural variant models is performed by a method selected from the group consisting of experimental methods, searching protein structure databases, homology modeling, molecular modeling, de novo protein folding, computational protein structure prediction, *ab initio* methods and combinations thereof.

105. (New) The method of claim 104, wherein the experimental methods include x-ray crystallography and NMR spectroscopy.

106. (New) The method of claim 13, wherein the step of determining 3-D protein structural variant models is performed by a combination of homology modeling and *ab initio* methods.

107. (New) The method of claim 15, wherein the step of determining 3-D protein structural variant models is performed by a method selected from the group consisting of experimental methods, searching protein structure databases, homology modeling, molecular modeling, de novo protein folding, computational protein structure prediction, *ab initio* methods and combinations thereof.

108. (New) The method of claim 107, wherein the experimental methods include x-ray crystallography and NMR spectroscopy.

109. (New) The method of claim 15, wherein the step of determining 3-D protein structural variant models is performed by a combination of homology modeling and *ab initio* methods.

110. (New) The method of claim 15, wherein the selected model structures represent structural variants derived from subjects who receive a specific treatment regimen.

111. (New) The method of claim 110, wherein a subject is a human.

112. (New) The method of claim 15, wherein the selected model structures represent structural variants derived from subjects who exhibit a particular clinical response to a given drug.

113. (New) The method of claim 112, wherein a subject is a human.

114. (New) The method of claim 15, wherein the selected model structures represent structural variants derived based on the duration of a particular drug treatment.

115. (New) The method of claim 15, wherein the target protein is an enzyme.

116. (New) The method of claim 115, wherein the enzyme is a protease or polymerase.

117. (New) The method of claim 116, wherein the polymerase is a reverse transcriptase.

118. (New) The method of claim 115, wherein the target protein is a protein expressed by an infectious agent.

119. (New) The method of claim 115, wherein the target protein is an enzyme expressed by an infectious agent.

120. (New) The method of claim 119, wherein the agent is a human immunodeficiency virus (HIV).

121. (New) The method of claim 15, wherein the target protein is a eukaryotic or prokaryotic protein.

122. (New) The method of claim 15, wherein the target protein is an animal protein, a plant protein or a protein from a pathogen.

123. (New) The method of claim 15, wherein the structural variant models are stored in a relational database, comprising:

3-D molecular coordinates for the structural variants;

a molecular graphics interface for 3-D molecular structure visualization; computer functionality for protein sequence and structural analyses; and

database searching tools.

124. (New) The method of claim 123, wherein the database further comprises one or more of observed clinical data associated with the genetic polymorphisms, subject medical history and subject history.

125. (New) The method of claim 1, wherein the target proteins are human proteins.

126. (New) The method of claim 1, wherein, based upon the intermolecular interactions between the 3-D models and drug candidates, drug candidates that preferentially interact with one of the target proteins are identified.

127. (New) The method of claim 1, wherein, based upon the intermolecular interactions between the 3-D models and drug candidates, an existing drug is modified so that it interacts with a plurality of the target proteins.

Please replace claims 1, 3, 15, 24, 25, 45, 48, 49, and 87 with the following amended claims:

1. (Amended) A computer-based method of drug design based on genetic polymorphisms, comprising:

identifying target proteins that are the product of a gene exhibiting genetic polymorphisms;

obtaining more than one amino acid sequence of the target proteins that are the product of a gene exhibiting genetic polymorphisms, wherein the sequences represent different genetic polymorphisms;

determining 3-dimensional (3-D) protein structural variant models for the target proteins that are the product of a gene exhibiting genetic polymorphisms; and

based upon the structures of the 3-D models of the target proteins that are the product of a gene exhibiting genetic polymorphisms, designing drug candidates, modifying existing drugs, identifying potential drug candidates or identifying modifications of existing drugs based on predicted intermolecular interactions of the drug candidates or modified drugs with the structural variants of the target proteins.

3. (Amended) The method of claim 2, wherein the binding interactions are determined by:

calculating the free energy of binding between the protein structural variant model and the docked molecule; and

decomposing the total free energy of binding based on the interacting residues in the protein active site.

4. (Amended) The method of claim 1, wherein:

after the protein structural variant models derived from a particular genetic polymorphism are generated, selected model structures are analyzed to determine common structural features that are conserved throughout the selected models, wherein

the conserved structural features are used as a basis for structure-based drug design studies.

15. (Amended) The method of claim 1, wherein:

after determining the 3-D protein structural variant models, the method comprises:

computationally docking drug molecules with the target protein models; and

energetically refining the docked complexes; and

wherein the candidate drugs are specific for a protein with a selected polymorphism or specifically interact with all proteins exhibiting a polymorphism.

24. (Twice Amended) The method of claim 15, wherein the structural variant models are stored in a relational database, comprising:

3-D molecular coordinates for the structural variants;

a molecular graphics interface for 3-D molecular structure visualization;

computer functionality for protein sequence and structural analysis; and

database searching tools.

F26 cont'd

25. (Amended) The method of claim 24, wherein the database further comprises observed clinical data associated with the genetic polymorphisms, subject medical history and subject history.

F27

45. (Amended) The method of claim 1, wherein the target protein is an enzyme.

F28

48. (Amended) The method of claim 45, wherein the target protein is a protein expressed by an infectious agent.

49. (Twice Amended) The method of claim 45, wherein the target protein is an enzyme expressed by an infectious agent.

F29

87. (Amended) The method of claim 12, wherein the selected subpopulation is a human subject subpopulation.

REMARKS

A check for the fee for a three month extension of time ($930) is enclosed. Any fees, including fees for additional claims and an extension of time, that may be due with this paper or with this application during its entire pendency may be charged to Deposit Account No. 50-1213. If a Petition for Extension of Time is required, this paper is to be considered such Petition.

Claims 1-25, 45-50, 66, 67, 87-90, and 95-127 are pending in the application. Claims 1, 3, 15, 24, 25, 45, 48, 49, and 87 are amended and claims 95-127 are added. Claim 1 is amended to render it clear that the 3-D structures of the proteins that are the product of a gene exhibiting genetic polymorphisms are determined. The amendment finds basis in claim 1 as originally filed; at page 4, lines 19-30, page 5, lines 3-13, page 21, lines 15-30, and pages 25-34, section 2, which begins as follows:

> After the amino acid sequences of target proteins are obtained via the means described in section 1, the 3-D structural models of the

sequences of native proteins or of the protein structural variants are then determined. They can be determined through experimental methods, such as x-ray crystallography and NMR, and from structure databases, such as the Protein Databank (PDB). Moreover, 3-D structural models can be determined by using any of a number of well known techniques for predicting protein structures from primary sequences (e.g. SYBYL (Tripos Associated, St. Louis, Mo.), *de novo* protein structure design programs (e.g. MODELER (MSI, Inc., San Diego, CA) and MOE (Chemical Computing Group, Montreal Canada) and *ab initio* methods, see, *e.g.*, U.S. Patent Nos. 5,331,573, 5,579,250 and 5,612,895), homology modeling, and *ab initio* computational analysis. Homology modeling, structure determination based upon x-ray crystallographic structures, and *ab initio* techniques and combinations of these methods are among those preferred herein.

Claims 3 and 4 are amended to add an inadvertently omitted comma for grammatical clarity. Claim 15 is amended to properly depend on amended claim 1; and claim 24 is amended to delete the conjunction "and" for grammatical clarity and to correct a claim dependency error to avoid redundancy. Original claim 24 depends on claim 13, which encompasses an embodiment "wherein the structural variant models are stored in a relational database." The amendment to claim 24 changes the dependency to claim 15, which has antecedent basis for "the structural variant models." The amendment to claim 24 also finds particular basis at page 34, lines 27-31, and at page 50, line 23 to page 52, line 13, which describe the storage of structural variant models into relational databases.

Claim 25 is amended to correct a claim dependency error and finds basis in original claim 24, which provides the antecedent basis for "the database." Claims 45 and 48 are amended to add the inadvertently omitted word "protein" and finds antecedent basis in claim 1, which recites "target proteins." Claim 49 is amended to add the inadvertently omitted article "an" for grammatical clarity and to add the inadvertently omitted word "protein." The amendment to claim 49 finds antecedent basis in claim 1 and in claim 45 as amended herein. Claim 87 is amended to correct a claim dependency error and to replace the word "patient" with the word —subject—. Original claim 87 depends on claim 1

which does not provide antecedent basis for "the selected subpopulation." Therefore, the amendment to claim 87 corrects this claim dependency by making claim 87 depend on claim 12, which provides the proper antecedent basis for "the selected subpopulation." The amendment to claim 87 also finds particular basis, for example, at page 2, lines 24-27; page 6, lines 1-2, 9, and 24; page 35, lines 14-25; and page 48, lines 9-14, which describes the subpopulations that are used in the methods of the instant application. No new matter has been added.

Dependent claim 95 is directed to the subject matter of cancelled claim 28, which was inadvertently omitted from the restriction requirement mailed on August 7, 2002, and which is considered to be a part of the elected subject matter in the Office Action mailed on December 3, 2002. Dependent claim 95 also finds basis, for example, at page 37, line 16, to page 39, line 13, of the specification and at Figure 3, which describe exportation of protein models into docking programs.

Dependent claims 96, 98, 100, 110, 112, and 114 find particular basis, for example, at page 35, line 25, to page 36, line 24, of the specification, which describes the criteria for selection of structural variant models. Dependent claims 97, 99, 111, and 113 find particular basis, for example, at page 2, lines 24-27; page 6 lines 1-2, 9, 24; page 35, lines 14-25; and page 48, lines 9-14, of the specification which describes subpopulations and subjects that are used in the methods of the instant claims. Dependent claims 115-122 find particular basis, for example, at page 24, lines 8-13, of the specification and the Examples, which describes target proteins. Dependent claims 101-109 find particular basis, for example, at page 25, line 17, to page 34, line 15, of the specification, which teaches ways of determining 3-D structures of the target proteins. Claims 123 and 124 find particular basis, for example, at page 34, lines 27-31; page 50, line 23 to page 52, line 13 of the specification and at Example 5, which describes storage of structural variant model data into

relational databases that are interfaced with molecular graphics for visualization, computer functionality for protein sequence and structure analysis, and database searching tools. Dependent claim 125 finds basis at page 24, lines 27-30, and page 82, lines 1-4, of the specification. Dependent claims 126 and 127 find basis at page 40, lines 5-9, of the specification and at claim 14. No new matter has been added.

The specification is amended to correct obvious typographical, formatting and grammatical errors. In particular, the amendment to the paragraph on page 34, lines 1-9, of the specification replaces the inadvertently labeled figure item "328" with —228— for clarity. The amendment finds basis in Figure 3. No new matter has been added.

The amendments to the drawings correct typographical errors.

Included as an attachment is a marked-up version of the specification paragraphs and claims that are amended.

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT

A Supplemental Information Disclosure Statement and Forms PTO-1449 (2 pages) making of record art cited in the Response and additional art are filed under separate cover on the same day herewith.

DRAWINGS

New formal drawings of corrected Figures 4, 6 and 7 are filed under separate cover. Copies of these formal drawings, marked-up copies of the original formal drawings of these figures and a copy of amended Figure 11 is attached to this response.

Figures 4, 6, and 7

In accordance with 37 C.F.R. §1.121(d), amendments to the drawings of Figures 4, 6, and 7 of the above-referenced application were submitted in a Preliminary Amendment dated March 21, 2001, for approval by the Examiner and not for the Office to amend the drawings. The amendments to these drawings were indicated in red ink per 37 C.F.R. §1.121(d).

Applicant respectfully requests that the Examiner reconsider and approve the amendments to Figures 4, 6, and 7. For the Examiner's convenience, a replacement copy of the hand-annotated pages of Figures 4, 6, 7 which show amendments to these figures in red ink is provided herewith; copies of the formal drawings thereof are attached.

The replacement formal drawings will be submitted under separate cover to the Drawings Review Branch.

Figure 11

Figure 11 was renumbered and a copy of these corrected drawings of Figure 11 were submitted under separate cover to the Drawings Review Branch on March 3, 2003. To be fully responsive, a copy of corrected Figure 11 is attached to this response.

THE REJECTION OF CLAIMS 1-25, 45-50, 66, 67, and 87-90 UNDER 35 U.S.C. §112, FIRST PARAGRAPH

Claims 1-25, 45-50, 66, 67, and 87-90 are rejected under 35 U.S.C. §112, first paragraph, as containing subject matter that was not described in the specification in such a way as to enable one of skill in the art to which it pertains, or with which it is most clearly connected, to make and/or use the claimed subject matter. Specifically, the Office Action states that generating 3-D protein structural variant models from the sequences would require undue experimentation because (a) there would be an unpredictable amount of experimentation required to determine the structure of a polypeptide from a polymorphic site by use of sequence data, (b) the specification does not present specific guidance to determine the structure of a polypeptide from sequence data, (c) the specification does not provide a working model of determination of the structure of a polypeptide from sequence data, (d) the nature of the claimed subject matter is complex, (e) the state of the art as represented by Sternberg *et al.* and Koehl *et al.* indicate that *ab initio* methods are unable to predict accurately, structures of complete polypeptides in the absence of knowledge of structures of polypeptides with similarity to the polypeptide of interest, (f) the

skill of those in the art of polypeptide structure modeling is high, (g) Sternberg *et al.* and Koehl *et al.* indicate that *ab initio* methods of structure prediction from polypeptide sequence information alone is not predicted to result in accurate structure of a complete polypeptide, and (h) the claims are broad in that they are drawn to modeling molecular structures in the absence of information other than sequence information. This rejection is respectfully traversed with respect to claims 1-25, 45-50, 66, 67, 87-90 and insofar as it applies to any of claims 95-127.

Summary of arguments below:

1) The scope of the claims is commensurate with the scope of enablement provided to one skilled in the art by the disclosure in the instant application; 2) the claims recite that the 3-D structures are determined for the target proteins; such determination can be effected by any of a variety of methods, including *ab initio* methods for generating 3-D structures of target proteins; and 3) notwithstanding 2), *ab initio* methods can be used to generate 3-D structures of proteins in the absence of additional information as exemplified in the application.

RELEVANT LAW

To satisfy the enablement requirement of 35 U.S.C § 112, first paragraph, the specification must teach one of skill in the art to make and use the invention without undue experimentation. *Atlas Powder Co. v. E.I. DuPont de Nemours*, 750 F.2d 1569, 224 USPQ 409 (1984). This requirement can be satisfied by providing sufficient disclosure, either through illustrative examples or terminology, to teach one of skill in the art how to make and how to use the claimed subject matter without undue experimentation. This clause does not require "a specific example of everything *within the scope* of a broad claim." *In re Anderson*, 176 USPQ 331, at 333 (CCPA 1973), emphasis in original. Rather, the requirements of 35 U.S.C. §112, first paragraph "can be fulfilled by

the use of illustrative examples **or** by broad terminology." *In re Marzocchi et al.*, 469 USPQ 367 (CCPA 1971)(emphasis added).

The inquiry with respect to scope of enablement under 35 U.S.C. § 112, first paragraph, is whether it would require **undue** experimentation to make and use the subject matter as claimed. A considerable amount of experimentation is permissible, particularly if it is routine experimentation. The amount of experimentation that is permissible depends upon a number of factors, which include: the quantity of experimentation necessary, the amount of direction or guidance presented, the presence or absence of working examples, the nature of the invention, the state of the prior art, the relative skill of those in the art, the predictability of the art, and the breadth of the claims (i.e. the "Forman factors"). *Ex parte Forman*, 230 USPQ 546 (Bd. Pat. App. & Int'f 1986); see also *In re Wands*, 8 USPQ2d 1400 (Fed. Cir. 1988).

PTO GUIDELINES

The standard for determining whether the specification meets the enablement requirement is whether it enables any person skilled in the art to make and use the claimed subject matter without **undue** experimentation. In re Wands, 858 F.2d 731, 737, 8 USPQ2d 1400 (Fed. Cir. 1999) (emphasis added). In determining whether any experimentation is "undue," the above-noted factors are to be considered.

As instructed in the published PTO guidelines, it is improper to conclude that a disclosure is not enabling based on an analysis of only one of the above factors while ignoring one or more of the others. The analysis must consider all the evidence related to each of the factors, and any conclusion of non-enablement must be based on the evidence as a whole. *Id.* 8 USPQ2d at 1404 & 1407.

The starting point in an evaluation of whether the enablement requirement is satisfied is an analysis of each claim to determine its scope. As set forth in the guidelines, all questions of enablement are evaluated against **the claimed**

subj ct matter. The focus of the inquiry is whether everything within the scope of the claim is enabled. With respect scope of enablement, the only relevant concern should be whether the scope of enablement provided to one skilled in the art by the disclosure is commensurate with the scope of protection sought by the claims. *In re Moore*, 439 F.2d 1232, 169 USPQ 236 (CCPA 1971). Once the scope of the claims is addressed, a determination must be made as to whether one skilled in the art is enabled to make and use the entire scope of the claimed invention without undue experimentation.

Analysis

Applying the above factors to the instant claims, applicant respectfully submits that as described in detail below, it would not require undue experimentation to practice the claimed methods.

1. **Breadth of the claims**

It is alleged in the Office Action that the claim methods requires an *ab initio* method to accurately predict the structure of a polymorphic polypeptide sequence in the absence of additional structural information (page 5, lines 20-22 of Office Action). Contrary to the Examiner's assertion, there is no reference in original claim 1 to *ab initio* methods for the determination of 3-D structures of the target proteins. In fact, as discussed below, the specification teaches that the step of determining 3-D structures can be performed by variety of methods (see, *e.g.*, pages 25-34) including, experimental methods, searches of protein structure databases, homology modeling, molecular modeling, de novo protein folding, computational protein structure prediction, *ab initio* methods and combinations of these methods. Therefore, it is improper for the Examiner to read into claim 1 a limitation that is not there. Claim 1 is a generic claim that encompasses any method for determining 3-D protein structures Therefore, it is improper for the Examiner to conclude that claim 1 and claims dependent thereon are limited to the species of *ab initio* methods for determining 3-D structures of target proteins.

It appears that the claims may have been misconstrued in the instant Office Action as requiring *ab initio* methods based on the phrase "generating 3-dimensional (3-D) protein structural variant models **from the sequences**" recited in claim 1. Applicant respectfully submits that the phrase does not refer to *ab initio* methods and, as described in the specification, is not be limited to *ab initio* methods (see, *e.g.*, pages 25-34, which set forth exemplary methods for determining 3-D structures); the sequences as set forth in the claims means from the proteins. As stated on page 25:

> After the amino acid sequences of target proteins are obtained . . . the 3-D structural models of the sequences of native proteins or of the protein structural variants are then determined. They can be determined through experimental methods, such as x-ray crystallography and NMR, and from structure databases, such as the Protein Databank (PDB). Moreover, 3-D structural models can be determined by using any of a number of well known techniques for predicting protein structures from primary sequences (e.g. SYBYL (Tripos Associated, St. Louis, Mo.), *de novo* protein structure design programs (e.g. MODELER (MSI, Inc., San Diego, CA) and MOE (Chemical Computing Group, Montreal Canada) and *ab initio* methods, see, *e.g.*, U.S. Patent Nos. 5,331,573, 5,579,250 and 5,612,895), homology modeling, and *ab initio* computational analysis. Homology modeling, structure determination based upon x-ray crystallographic structures, and *ab initio* techniques and combinations of these methods are among those preferred herein.

As noted above, the claims must be properly construed in order to perform an enablement analysis as mandated by the PTO guidelines and caselaw. Because the Examiner has <u>improperly construed</u> the claims as requiring solely *ab initio* methods for generating 3-D structures of target proteins, it is respectfully submitted that the rejection of the instant claims as not being enabled by the specification cannot be valid since the basis for it set forth in the instant Office Action reflects an improper characterization of the subject matter that is claimed.

To effect clarity, applicant has amended claim 1 to recite that the 3-D structures of the target proteins are determined, so that claim 1 will not be misconstrued as requiring *ab initio* methods for 3-D structures of target proteins

that are the product of a gene exhibiting genetic polymorphisms. Claim 1 and claims dependent thereon are discussed below.

The claims

1. (Amended) A computer-based method of drug design based on genetic polymorphisms, comprising:

identifying target proteins that are the product of a gene exhibiting genetic polymorphisms;

obtaining more than one amino acid sequences of the target proteins that are the product of a gene exhibiting genetic polymorphisms, wherein the sequences represent different genetic polymorphisms;

determining 3-dimensional (3-D) protein structural variant models for the target proteins that are the product of a gene exhibiting genetic polymorphisms; and

based upon the structures of the 3-D models of the target proteins that are the product of a gene exhibiting genetic polymorphisms, designing drug candidates, modifying existing drugs, identifying potential drug candidates or identifying modifications of existing drugs based on predicted intermolecular interactions of the drug candidates or modified drugs with the structural variants of the target proteins.

Hence the claim is directed to computer-based methods of drug design in which 3-D models of structural variant target proteins that are the product of a gene exhibiting genetic polymorphisms are determined and used for a variety of purposes, including designing drug candidates, modifying existing drugs, identifying potential drug candidates or identifying modifications of existing drugs based on predicted intermolecular interactions of the drug candidates or modified drugs with the structural variants of the target proteins. Claims 2-25, 45-50, 66, 67, and 87-90 are dependent on claim 1 and further define the structure based drug design, target proteins, selection of model structures, determination of binding interactions, analysis of structure of target proteins to determine common structural features, and storage of structures in a database. Claims 95-125 are dependent on claim 1 and further describe computational docking, determinatin of 3-D structures of target proteins, and selection of model structures, and protein source.

2. Th amount of direction and guidance presented, teachings of the specification

Each step of the computer-based drug design method is taught in the instant specification.

a) Identifying target proteins that are the product of a gene exhibiting genetic polymorphisms

The specification teaches that target proteins are identified as being the product of a gene exhibiting genetic polymorphisms (page 24, lines 3-27). A target protein is a protein, polypeptide, or oligopeptide that includes, but is not limited to, receptors, enzymes, hormones, prions, or any such compound with which drugs or other ligands, such as small molecules, peptide agonists, peptide antagonists, other proteins, nucleic acids and other biomacromolecules, interact to bring about a biological response.

b) Obtaining amino acid sequences of target proteins

The specification teaches how to obtain one or more amino acid sequences for a target protein that is the product of a gene exhibiting genetic polymorphisms (See e.g. page 24, line 3 to page 25, line 16). For example, amino acid sequences can be obtained by the isolation and analysis of the gene or gene products in samples taken from pathogens (e.g. protozoans, viruses, including DNA and retroviruses, and bacteria), parasites, plants, animals, and humans. Amino acid sequences can also be identified and obtained from protein isolation and sequences methods, from data banks and databases (e.g. GenBank, Swiss Prot, PIR), from blood and tissue banks, and from publications.

c) Determining 3-D protein structures of target proteins that are the product of a gene exhibiting genetic polymorphisms.

Contrary to the Examiner's assertion, claim 1 and claims dependent thereon do not solely require *ab initio* methods for determining 3-D structures of target proteins that are the product of a gene exhibiting genetic polymorphisms. Additionally, as taught in the specification, *ab initio* methods can be used to determine 3-D protein structures of these target proteins. As discussed above

(see, *e.g.*, pages 25-34), after the amino acid sequences of target proteins are obtained, 3-D structural models of the sequences of native proteins or of the protein structural variants are then determined. As stated in the specification (page 25-26), they can be determined by a variety of methods:

> They can be determined through experimental methods, such as x-ray crystallography and NMR, and from structure databases, such as the Protein Databank (PDB). Moreover, 3-D structural models can be determined by using any of a number of well known techniques for predicting protein structures from primary sequences (e.g. SYBYL (Tripos Associated, St. Louis, Mo.), *de novo* protein structure design programs (e.g. MODELER (MSI, Inc., San Diego, CA) and MOE (Chemical Computing Group, Montreal Canada) and *ab initio* methods, see, *e.g.*, U.S. Patent Nos. 5,331,573, 5,579,250 and 5,612,895), homology modeling, and *ab initio* computational analysis. Homology modeling, structure determination based upon x-ray crystallographic structures, and *ab initio* techniques and combinations of these methods are among those preferred herein.

Pages 26-34 then describe some of these methods in detail. For example, page 34 states:

> The crystal structure of any protein can be determined empirically and the resulting coordinates used as the basis for determining [*sic*] structures of variants. Such structures are often known (see, *e.g.*, Kohlstaedt *et al.* (1992) *Science 256*:1773-1790 for a crystal structure of HIV-1 RT bound to a ligand).

Furthermore, there is nothing in the art nor in the specification nor cited by the Examiner that demonstrates that *ab initio* methods alone cannot be used to generate 3-D structures of polypeptides.

***Ab Initio* Methods**

Contrary to statements in the Office Action, *ab initio* methods can be used to generate 3-D protein structures, and such methods were known before the effective filing date of the application. For example, the specification teaches that *ab initio* methods, such as those taught in U.S. Patent Nos. 5,331,573; 5,579,250; and 5,612,895, cited on page 28, lines 17-22, of the instant application and incorporated by reference into the instant application can be used to generate 3-D protein structures. The specification teaches that these

methods involve simulating a real-size primary structure of a polypeptide in a solvent box, i.e., an aqueous environment; shrinking the size of the peptide isobarically and isothermally; and expanding the peptide to its real size in selected time periods, while measuring the energy state and coordinates, i.e., the bonds, angles and torsions of the expanding molecule (page 28, lines 22-27). The specification teaches that as the peptide expands to its full size, it assumes a stable tertiary structure and that, in most cases, this tertiary structure will be either the most probable structure (i.e., it will represent a global minimum for the structure) or one of the most probable structures (page 28, lines 27-31). The specification also teaches that once a model is built, it can be refined using energy minimization, molecular dynamics calculation or simulated annealing (page 29, lines 4-6). The steric and energetic quality of the structural models can be evaluated by analyzing the structural attributes of the model, such as phi and psi angles or by analyzing the energetics of the model, such as by calculating energy per residue or strain energy (page 29, lines 6-10).

These methods for *ab initio* generation of 3-D structures are disclosed and claimed in the above-noted patents. U.S. Patent Nos. 5,331,573, 5,579,250 and 5,612,895 each to Balaji *et al.* describe and claim methods of rational drug design that include simulating polypeptides in a way that predicts the most probable secondary and/or tertiary structures of polypeptides without any presumptions as to the conformation of the underlying primary or secondary structure. Once the most probable conformation of the polypeptide of interest is selected, analogs are designed and synthesized and evaluated for bioactivity. Additionally, peptidomimetics based on the conformation of the synthesized analogs are designed. Thus, the 3-D protein models generated by the Balaji *et al. ab initio* methods are useful in drug design methods.

In fact, the Balaji *et al. ab initio* methods were successfully used to predict the 3-D structures of endothelin-1 and to design bioactive analogs thereof (U.S. Patent No. 5,736,509). In particular, the drug design methods

include the steps of (a) simulating the most probable conformations of endothelin-1 and selecting the most probable conformation from among the simulated conformations; (b) designing and synthesizing cyclic peptides that mimic selected surface features of the three-dimensional structure of endothelin-1; and (c) evaluating the bioactivity of the cyclic peptides (U.S. Patent No. 5,736,509). Cyclic peptides claimed in U.S. Patent No. 5,736,509 mimic the surface of endothelin-1 and act as **antagonists and otherwise modulate the activity of endothelin**. Thus, the Balaji *ab initio* methods are adequate for generating 3-D polypeptide structures for the design of bioactive molecules and drug analogs. The Examiner has provided no reasons why the Balaji *et al. ab initio* methods cannot be used to generate 3-D structures of variant models for use in the instantly claimed methods of drug design based on genetic polymorphisms described in the instant application.

Other Methods

Notwithstanding the fact that *ab initio* methods are adequate, the specification does not limit generating 3-D structures of target proteins to *ab initio* methods and teaches **a variety of ways** to generate 3-D structures of all or a portion of protein structural variant models from protein sequences (See e.g. page 21, lines 15-30, and page 25, line 17, to page 34, line 15). The step of determining 3-D protein structural variant models in claims 1-25, 45-50, 66, 67, and 87-90 and claims 95-125 includes any method by which such model can be constructed. For example, as noted several times above, the specification teaches that 3-D protein structural variant models can be determined from experimental methods, such as x-ray crystallography and NMR, from structure databases, such as the Protein Databank (PDB), from well known techniques for predicting protein structures from primary sequences (e.g. SYBYL (Tripos Associated, St. Louis, Mo.), *de novo* protein structure design programs (e.g. MODELER (MSI, Inc., San Diego, CA), from MOE (Chemical Computing Group, Montreal Canada), from *ab initio* methods, (see, *e.g.*, U.S. Patent Nos.

5,331,573, 5,579,250 and 5,612,895), from homology modeling, and from *ab initio* computational analysis (page 25, lines 17-30) and combinations thereof.

X-ray crystallographic data

The specification provides that the coordinates of the crystal structure of proteins can be used as the basis for determining structures of protein variants (page 34, lines 10-15). Therefore, the specification teaches a skilled artisan how to use determine 3-D protein structure via x-ray crystallographic data.

Homology Modeling

The specification teaches that one of skill in the art can determine 3-D protein structural variant models by, for example, homology modeling in which proteins of unknown structure can be constructed using composite parts of related proteins with known structures (i.e. reference proteins; page 26, line 3, to page 28, line 15). A 3-D model can be built by incorporating the structural attributes of the reference protein together with the sequence of the target protein (page 26, lines 12-14). Coordinates from the protein backbone of the reference proteins can, be used to construct the backbone framework for the target protein structure, and side chains can be constructed, for example, by using side chain coordinates from the reference proteins, by searching a database to obtain side chain conformations that fit in with the existing structural framework, or by generating side chains *ab initio* to establish energetically favorable side chain conformations (page 27, lines 9-16). The non-conserved regions of the protein of unknown structure can be constructed by using, for example, database searching to identify other proteins with similar variable regions or by fitting the target sequence to a peptide backbone generated by varying phi and psi angles (*e.g.*, by calculating Ramachandran or Balasubramanian plots (page 27, lines 17-18 and lines 26-28). Sequence homology studies can be carried out using sequence alignment algorithms well known in the art (page 27, lines 22-25). Homology modeling can be used in conjunction with *ab initio* loop prediction or *ab initio* secondary

structure prediction methods (Figures 2-3 and page 29, lines 22-28) or conjunction with the *ab initio* methods described below to generate 3-D protein structures. Therefore, the specification teaches a skilled artisan how to use determine 3-D protein structure via homology modeling.

Combination of Homology Modeling and *ab initio* methods

The specification also teaches that a combination of homology modeling and *ab initio* methods can be used to determine structural variant models (page 29, lines 22-28). In particular, protein sequence information that is derived based on the genetic polymorphisms is used to assign the protein to a protein superfamily in order to identify related proteins to be used as templates to construct a 3-D model of the protein. If the superfamily is not known, sequence analysis or structural similarity searched can be performed to identify related proteins for use as templates in homology modeling studies (page 29, lines 16-21). *Ab initio* loop prediction or *ab initio* secondary structure generation techniques are then used to complete the model (as shown in Figures 1-3). The structural variant model can be energetically refined by performing molecular mechanics calculations, for example, using an ECEPP type force field or through molecular dynamics simulations, for example, using a modified AMBER type force field. If necessary, the structures can be dynamically refined, for example, by using a simulated annealing protocol (e.g., 100 ps equilibration, 500 ps dynamics, up to 1000°K, 1 fs data collection). For quality control, the protein structural characteristics, for example, stereochemistry e.g., phi/psi and side chain angles), energetics (e.g., strain energy), packing profile (e.g., packing factor per residue) and hydrophobic packing are evaluated and required to meet acceptable criteria before the structures are used in further studies (e.g. drug design methods described below) or input into a structural polymorphism database (page 31, lines 1-6).

Therefore, not only are *ab initio* methods adequate, but as taught in the specification, the step of "determining 3-D protein structural variant model" in

claims 1-25, 45-50, 66, 67, 97-90 can be done by a variety of methods including experimental methods (e.g., x-ray crystallography and NMR), from a protein structure database (e.g.the PDB), homology modeling, *de novo* protein folding algorithms and methodologies, other *ab initio* methods, and other computational protein structure prediction methods.

d) Drug Design based on the 3-D Structure of Target Proteins that are the product of a gene exhibiting genetic polymorphisms

The specification also teaches how to use 3-D structural variant models to design drugs that interact with 3-D structures or portions of 3-D structures of protein structural variants. For example, structure-based drug discovery methodologies such as computational screening or docking (e.g. DOCK and AUTODOCK) can be used to design biologically active compounds that interact with the 3-D structures of the target proteins or portions of target proteins (page 21, lines 15-30, and page 23, lines 10-20 and page 36, lines 25-30). Using these methods, drug designers can identify and computationally rank the various potential clinical drug candidates for maximum efficacy, thereby performing drug discovery *in silico* (page 23, lines 16-18). The drug design methods of the instant application are based on the intermolecular interactions between drug candidate or modified drugs and the selected structural variants that are predicted by computationally docking drug molecules with the target protein models; energetically refining the docked complexes; determining the binding interactions between the drug or potential new drug candidate molecules and the models by calculating the free energy of binding of the docked complexes and decomposing the total free energy of binding based on interacting residues in the protein active site or sites deemed important for protein activity (page 37, lines 7-15).

The specification teaches that structure-based drug design experiments, such as computational screening or docking studies, calculation of binding energies or analysis of steric, electrostatic, or hydrophobic properties of the resulting structural variant models, can be performed on selected structural variant models to aid in the understanding of observed biological activities or to determine new potential drug candidates to bind to the particular target (page 37, lines 20-26). For example, databases of molecules can be screened for molecules that complement the structural variant models through, for example, steric, hydrophobic, or electrostatic interactions. Additionally, solvation parameters can be factored in (page 37, lines 27-31). New potential drug candidates can be designed by identifying potential small molecule drugs that bind to a particular structural variant, by modifying the functional groups of existing drugs *in silico*, or by using methods of *de novo* ligand design well known in the art (e.g. GRID (Oxford University, Oxford, UK); CAVEAT (UC Berkeley, Ca), LEGEND (Molecular Simulations, Inc., San Diego, CA); LUDI (Molecular Simulations, Inc., San Diego, CA); HOOK (Molecular Simulations, Inc., San Diego, CA); CLIX (CSIRO, Australia); GROW (Upjohn Laboratories, Kalamazoo); others including HINT, LUDI, NEWLEAD, HOOK, PRO-LIGAND and CONCERTS (see, M. Murcko, "An Introduction to De Novo Ligand Design" in Practical Application of Computer-Aided Drug Design, Charifson, Ed., Marcel Dekker, NY, pp 305-354; page 38, line 26, to page 39, line 3 of specification). The docked complexes, if needed, can be further refined energetically using, for example, molecular mechanics, molecular dynamics, and simulated annealing techniques, to optimize geometries within the binding site and to select the best structure from a set of possible structures. The free energy of binding of the docked complex is calculated, and analyses of the binding energies are used to identify drug candidates (line 39, lines 15-20). Since the target protein

structural variants are based on genetic polymorphisms, the identified drug candidates that interact with these structures can be used to target specific populations or subpopulations, to target a specific genetic polymorphism, or to target a maximum number of genetic polymorphisms.

3. Presence of working examples

The specification provides working examples of a variety of ways of determining 3-D structural variant models of target proteins that are the product of a gene exhibiting genetic polymorphisms.

The specification provides working examples of the determination of protein structure using x-ray crystallographic data and molecular modeling. In Example 1, the structure of NS3-peptide complexes is determined using sequence data in conjunction with crystallographic data of NS3/NS4A peptides and molecular modeling using Monte Carlo simulations and ECEPP/3 force field. In Example 1, 3-D models of mutant forms of HCV protease with different inhibitors are determined and used to assess binding correlations between experimental results and computational results. In particular, binding correlations of NS3/NS4A-peptide complexes were studied and used to validate the 3-D models of the NS3/NS4A-peptide complexes.

NS3 is an approximately 68 kDa protein that is involved in HCV replication and is thus a target for anti-viral drugs. NS3 protease is a serine protease that is responsible for proteolysis of the polypeptide (polyprotein) at the NS3/NS4a, NS4a/NS4b, NS4b/NS5a and NS5a/NS5b junctions responsible for generating four viral proteins during viral replication. NS3 protease is inhibited by N-terminal cleavage products of substrate peptides.

Active NS3 forms a heterodimer with a polypeptide cofactor NS4A. The crystal structure of NS3 with and without the NS4A cofactor is known in the art. Example 1 describes binding studies of NS3/NS4A complexes with two peptide inhibitors. The binding studies were conducted *in silico* and compared

to the experimental results of Ingallinella *et al.* ((1998) *Biochemistry* *37*:8906-8914).

Generation of 3-D structure of NS3/NS4A-peptide complex

The crystal structure of NS3/NS4A was regularized using molecular mechanics. Peptides were placed into the NS3 binding site by analogy with other serine proteases. The peptides studied were:

Sequence *	IC^{50}, nM	SEQ ID
Ac-Asp^1-D-Glu^2-Leu^3-Ile^4-Cha^5-Cys^6-COO-	15	1
Ac-Asp^1-L-Glu^2-Leu^3-Ile^4-Cha^5-Cys^6-COO-	60	2

* Cha = β-cyclohexylalanine

Monte Carlo (MC) simulations were performed on the NS3/NS4A-peptide complexes using ECEPP/3 force field. The sampling method was a biased probability Monte Carlo with random change of one variable at a time. A Metropolis acceptance criterion was applied after energy minimization (quasi-Newton, up to 1000 steps). Simulations were performed at a temperature of 1000° K. In the peptide, translation/rotation and all torsions were included in the simulation. Protein side-chain χ angles of residues that have at least one atom within 7.0 Å of any atom of the peptide were included. The energy function used in the MC simulations included ECEPP/3 terms for energy *in vacuo* (VDW, H-bond, electrostatic and torsion potentials); distance dependent electrostatics with e = 4.0; and surface energy with atomic solvation parameters. The total energies of the complexes were calculated including contributions from: ECEPP/3 VDW, H-bond, S-S bond and torsion terms; exact-boundary electrostatic energy with e=8.0; and side-chain entropies. Hydrophobic free energies were estimated as sA, where A is accessible surface area and s is a tension constant of 0.03 kcal/mol$Å^2$.

The Monte Carlo (MC) simulations of the NS3/NS4A-peptide complexes were performed with multiple, relatively short MC runs (2000-5000 generated structures). New docking cycles were started from the lowest-energy or other interesting structures found in previous runs. Structures saved during various MC runs were sorted by total energies and RMSD, and compressed into a cumulative conformational stack.

Binding energies were calculated for representative structures of each complex thus obtained using the equation:

$$E_{bind} = E_o + E_{compl} - E_{pept} - E_{prot},$$

where E_{compl} is the energy of the NS3/NS4A-peptide complex, E_{pept} & E_{prot} are separate energies of the peptide and NS3/NS4A protein, respectively, and E_o is an adjustable constant. The binding energy function included: exact-boundary electrostatic contributions; side-chain entropy; and surface tension hydrophobic terms. ECEPP/3 hydrogen-bonding terms were included with a weight of 0.5.

Validation of Models

Modifications were made to the NS3/NS4A-peptide complex models, and the binding energies of the modified complexes were correlated with those expected from experimental IC_{50} values. Changes in calculated binding energies upon modifications, ΔE_{bind} (calc), were compared to the values expected from ratios of inhibitory potencies, ΔE_{bind}(exp).

$$\Delta E_{bind}(exp) = RT\, ln(IC_{50}^{mod}/IC_{50}^{o}),$$

where IC_{50}^{o} and IC_{50}^{mod} are inhibitory potencies of the parent and modified compounds.

Predicted changes in binding energy upon modification of the protein and peptides correlate reasonably well with the changes expected from IC^{50} ratios (see Figure 4). For example, standard deviations of ΔE_{bind}(calc) - ΔE_{bind}(exp) were 0.8 and 1.6 kcal/mol for Models 1 and 2, respectively, with correlation coefficients of 0.62. After the largest outlier was removed from each dataset, correlations improved to 0.81 and 0.76, respectively.

Thus, the fact that the *in silico* studies correlated well with experimental data validated these models for use in structure-based drug design methods and for other methods claimed in the instant application. Moreover, the correlation of the computational data with experimental data demonstrates that the skilled artisan can use the protocols taught in Example 1 to generate 3-D structural variant models.

In Example 2, thirteen active compounds were docked to the 3-D structure of TNF receptor type I protein. The binding studies conducted in silico were consistent with the corresponding experimental values, with a correlation coefficient of 0.966 (Figure 4), thus validating the protein model. In addition to these docking studies, graphics-guided redesign of the active compounds yielded new compounds that had a high affinity (such as a K_i in low nanomolar range) for the target TNF receptor type I protein.

In Example 3, computational binding studies were conducted on HIV protease-inhibitor complexes. The initial structure of the proteins were obtained from the Protein Data Bank. The proteins were complexed with FDA approved inhibitors (i.e. saquinavir, nelfinavir, indinavir, amprenavir, and ritonavir), and the complexes were optimized using Monte Carlo simulations and ECEPP/3 force field using the protocol described in Example 1. Free energy of binding analyses showed a correlation of 0.9 between computationally calculated and experimentally determined binding energy data for HIV protease-ritonavir and HIV protease-amprenavir complexes (Table 3), thus validating the model generation protocol of Example 1. Additionally, in Example 3, amino acid sequences of HIV protease were aligned, and the frequency of mutation was determined at each amino acid position to identify invariant and/or highly conserved regions of the protease and to map these regions onto the 3-D structure for protease.

4. Natur of the claimed subject matter

The claimed subject matter is directed to a computer-based method of drug design in which 3-D protein structural variant models of target proteins that are the product of a gene exhibiting genetic polymorphisms are determined and based upon the interaction of the 3-D models of the target protein used for a variety of drug design purposes.

As noted above, the 3-D protein structural variant models of these target proteins can be determined by experimental methods, searching protein structure databases, homology modeling, molecular modeling, de novo protein folding, computational protein structure prediction, *ab initio* methods and combinations thereof.

As taught in the specification, the drug design methods of the instant application are based on the intermolecular interactions between drug candidates or modified drugs and the selected structural variants that are predicted by computationally docking drug molecules with the target protein models; energetically refining the docked complexes; and determining the binding interactions between the drug or potential new drug candidate molecules and the models by calculating the free energy of binding of the docked complexes and decomposing the total free energy of binding based on interacting residues in the protein active site or sites deemed important for protein activity (page 37, lines 7-15). As also taught in the specification, structure-based drug discovery methodologies such as computational screening or docking programs such as, but not limited to, DOCK (available from University of Ca, San Francisco); AUTODOCK (available from Scripps Research Institute, La Jolla); GRID (Oxford University, Oxford, UK); CAVEAT (UC Berkeley, Ca); LEGEND (Molecular Simulations, Inc., San Diego, CA); LUDI (Molecular Simulations, Inc., San Diego, CA); HOOK (Molecular Simulations, Inc., San Diego, CA); CLIX (CSIRO, Australia); GROW (Upjohn Laboratories, Kalamazoo); others including HINT, LUDI, NEWLEAD, HOOK, PRO-LIGAND and

CONCERTS (see, M. Murcko, "An Introduction to De Novo Ligand Design" in Practical Application of Computer-Aided Drug Design, Charifson, Ed., Marcel Dekker, NY, pp 305-354) can be used to screen, design, or modify biologically active compounds that interact with the 3-D structures of the target proteins (page 23, lines 10-20 and page 36, lines 25-30).

As also taught in the specification, the 3-D structures of the target proteins that are the product of a gene exhibiting polymorphisms can be compared to each other and analyzed to determine common structural features that are conserved throughout the selected models (page 5, lines 25-31). Once the common structural features are determined, these features can be used as the basis for drug design using the methods described above.

As also taught in the specification, the 3-D structures of the target proteins that are the product of a gene exhibiting polymorphisms can be selected based on variety of criteria, including, but not limited to, 3-D structures of proteins that result from the most commonly occurring genetic polymorphisms, that result from genetic polymorphisms found in a selected subpopulation, that are derived from subjects who receive a specific treatment regimen, that are derived from subjects who exhibit a particular clinical response to a given drug, or that are derived based on the duration of a particular drug treatment (page 35, line 25, to page 36, line 24). The 3-D structures of target proteins that are selected based on the criteria listed above are used as the basis for drug design using the methods described above, such that the drug design methods can be used to target the most commonly occurring polymorphisms, specific subpopulations and subjects, or particular treatment regimens.

5. Level of skill

The level of skill in this art of polypeptide structure and modeling is recognized to be high. In addition, the numerous articles and patents that are of record in this application that are authored by those of a high level of skill for an

audience of a high level of skill further evidences the high level of skill in this art. For example, Balaji *et al.* disclose *ab initio* methods that can be used to generate polypeptide structures without additional information on structure (U.S. Patent Nos. 5,331,573; 5,579,250; and 5,612,895). Dudek *et al.* discloses *ab initio* loop predictions and molecular modeling that can be used to generate polypeptide structures (J. Comp. Chem. (1998) 19:548-573). Balasubramanian discloses Ramachandran plots that can be used to generate polypeptide structures (Nature (1974) 266:856-857). Weiner and Ramnarayan disclose molecular modeling that can be used to generate polypeptide structures (J. Comp. Chem. (1986) 7:230-252; J. Chem. Phys. (1990) 92:7057-7076). Nanni *et al.* and Kroeger *et al.* disclose 3-D protein structures of HIV reverse transcriptase and protease (*Perspectives in Drug Discovery and Design 1*:129-150 (1993); *Protein Eng. 10*:1379-1383 (1997)). Love *et al.* and Yan *et al.* disclose the crystal structure of NS3 protein without the NS4A cofactor (*Cell 87*:331-342 (1996) and *Protein Sci. 7*:837-847 (1998)). Since the level of skill in the art is high, the specification does not require a high level of teaching in order to enable one of skill in the art to generate 3-D protein structures.

6. State of the prior art:

At the time of the effective filing date of this application and before, the skilled artisan knew various methods of determining 3-D structures. Further, there is a large body of literature directed to determining 3-D structures of proteins from experimental and computer-based methods.

The articles cited in the specification, of record in this application, describe various methods of determining 3-D protein structures. The following represents some exemplary articles that evidence the knowledge of those of skill in the art at the time of filing of the instant application:

Balaji *et al.*, "Method of Rational Drug Design Based on Ab Initio Computer Simulation of Conformational Features of Peptides," U.S. Patent No. 5,612, 895 (March 18, 1997), which describes *ab initio* prediction of the tertiary structure of polypeptides without any presumption as to the conformation of the underlying primary or secondary structure, as discussed above. The Balaji *et al.* method involves computer simulation of polypeptides by simulating a real-size primary structure in an aqueous environment, shrinking the size of the polypeptide isobarically and isothermally, and expanding the simulated polypeptide to its real size in selected time periods. "Balaji plots" are used to identify those portions of the predicted peptide structure that are most flexible or rigid. The Balaji *et al.* methods are used to generate polypeptide models that are used to design bioactive drug analogs (See U.S. Patent No. 5,736,509 (April 7, 1998); See also Balaji *et al.* U.S. Patent No. 5,331,573 (July 19, 1994) and U.S. Patent No. 5,579,250 (November 6, 1994));

Dudek *et al.* "Protein Prediction Using a Combination of Sequence Homology and Global Energy Minimization: II. Energy Functions," (1998) J. Comput. Chemistry 19:548-573, which describes the application of *ab initio* methods using global energy minimization to reconstruct surface loops that were stripped from the protein crystal structures of proteins such as avian pancreatic polypeptide, crambin, bptl, erabutoxin B, immunoglobulin domain, ribonuclease A, and lysozyme. The surface loops that were reconstructed using *ab initio* methods were in good agreement with the protein crystal structure with RMSD < 1 with a weighting factor of 1 (Table IX).

Abagyan and Totrov, "Ab Initio Folding of Peptides by the Optimal-Bias Monte Carlo Minimization Procedure," (May 1, 1999) 151:402-421, which describes the structure prediction of BBA5 protein-like folded peptide using *ab initio* methods employing a Monte Carlo minimization-based global optimization algorithm (i.e. OBPMC) in conjunction with a set of biased random moves. Using the OBPMC method, Abagyan and Totrov generated low energy

conformations of the BBA5 protein and compared it to the experimental structure of BBA1 (an analog of BBA5 with the synthetic residue Fen at the 3 position) and observed similarity between the *ab initio* generated structure and the experimental structure of the analog (i.e. RMSD 3.8).

Osguthorpe, "Improved Ab Initio Predictions with a Simplified, Flexible Geometry Model," Proteins: Structure, Function, and Genetics Suppl 3 (November 8, 1999) 186-193, which describes the *ab initio* prediction of the tertiary structure of T56, (a target that was predicted at the CASP3 meeting described by Sternberg *et al.* and Koehl *et al.*) and which demonstrates that for new folds, *ab initio* methods are as good as other methods in getting an approximation to the native structure;

Westhead and Thornton "Protein structure prediction," Curr Opin in Biotechnology (1998) 9:383-389, which describes improvements in comparative modeling studies;

Eisenhaber *et al.* "Protein structure prediction: recognition of primary, secondary, and tertiary structural features from amino acid sequence," Critical Rev. in Biochem and Mol. Biol. (1995) 30:1-94, which describes prediction of protein structure from amino acid sequence computational methods, threading methods, sequence alignment methods, and homology modeling;

Jones, "Successful *ab initio* prediction of the tertiary structure of NK-Lysin using multiple sequences and recognized supersecondary structural motifs," Proteins: Structure, function, and Genetics, Suppl 1 (1997) 185-191, which describes tertiary protein structure prediction based on the assembly of recognized structural fragments taken from highly resolved protein structures;

Samudrala *et al.*, "Ab initio protein structure prediction using a combined hierarchical approach," Proteins: Structure, function, and Genetics Suppl 3 (November 8, 1999) 194-198, which describes the prediction of 3-D structures for 13 proteins using lattice based scoring function in conjunction with distance geometry methods;

Dunbrack *et al.* "Meeting review: the Second Meeting on the Critical Assessment of Techniques for Protein Structure Prediction (CASP2), Asilomar, California, December 13-16, 1996," Folding and Design (1997) R27-R42, which reviews the comparative modelling, fold recognition, and *ab initio* structure prediction presented at CASP2; and

de Dios *et al.* "Secondary and Tertiary Structural Effects on Protein NMR Chemical Shifts: An ab Initio Approach," Science (1993) 260:1491-1496, which describes the use of NMR to refine protein structures;

These articles and patents are representative of the numerous ways to determine 3-D protein structures. The articles range from pure experimental methods (such as NMR spectroscopy) to pure computational methods (such as *ab initio* methods of Balaji *et al.*). Therefore, the knowledge of those of skill in the art is extensive and methods for determining 3-D protein structures is well known.

7. Predictability

As discussed above, the individual steps in the methods are described and exemplified in the application and are based upon known methods for determining 3-D structures and predicting interactions of the 3-D structures with compounds. Since each step of the method can be practiced using technology and methods available to those of skill in the art, the combination of steps of the methods will work as claimed.

CONCLUSION

In light of the breadth of the claims, the teachings, guidance, and working examples in the specification, the high level and extensive knowledge of skill of those in this art, it would not require undue experimentation for a person of skill in the art to generate 3-D protein structures for use in the instant method claims. As disclosed in the specification and in the cited art, the skilled artisan can use experimental methods, computational methods, homology modeling, and/or *ab initio* methods to generate 3-D protein structures. Therefore, the

specification is enabling for making and using the full scope of the claimed subject matter.

Rebuttal to comments by the Examiner

1. Rebuttal to Examiner's state of the art arguments

The Examiner alleges that the instant methods require *ab initio* methods to predict accurately the structure of a polymorphic peptide sequence in the absence of additional structure and that the best *ab initio* methods were unable to predict protein structure accurately in the absence of knowledge of structures of polypeptides with similarity to the polypeptide of interest as evidenced by Sternberg *et al.* (Curr Opin in Struc. Biol. (1999) 9:368-373) and Koehl *et al.* (Nature Struc. Biol. (1999) 6:108-111).

The instant claims are directed to computer-based methods of drug design based on genetic polymorphisms. Candidate drugs are designed or modified based on their interactions with the 3-D structures of target proteins that are the product of a gene exhibiting genetic polymorphisms. Experimental methods (e.g. NMR and x-ray crystallography), computational methods, homology modeling, de novo prediction, *ab initio* methods, and other methods and combination thereof can be used to determine 3-D protein structures of the target proteins that the product of a gene exhibiting genetic polymorphisms, as noted above. Thus, the claims do not require *ab initio* methods for generation of 3-D structures.

Notwithstanding the variety of methods that can be used to determine 3-D structures of these target proteins, *ab initio* methods are adequate for generating 3-D protein structures of target proteins, as demonstrated above.

Therefore, Examiner has improperly construed certain references to support a conclusion that *ab initio* methods are inadequate. For example, Sternberg *et al.* and Koehl *et al.* are review articles from the third comparative assessment of techniques of protein structure prediction (CASP3) meeting held during 1998. These articles describe three approaches to protein structure

prediction, namely comparative modeling, fold recognition, and *ab initio* prediction. The Examiner alleges that both of these references disclose that *ab initio* algorithms had relatively poor ability to predict structure and predicted the positions of only about half of the residues in the sequence in the absence of knowledge of the structures of other similar polypeptides (Sternberg *et al.* Figure 2 and Koehl *et al.* Table 1).

Figure 2 in Sternberg *et al.* shows the best predictions of protein targets in the fold recognition and *ab initio* sections. Contrary to the Examiner's assertion that *ab initio* algorithms had relatively poor ability to predict structure, Sternberg *et al.* discloses that in the *ab initio* section, algorithms were available for generating good models for several targets from sequence without employing fold recognition (Sternberg *et al.* page 371, column 2, lines 31-34). Figure 2 also shows that for two targets, *ab initio* methods yielded better models than fold recognition methods that used a template fold that was a distant homologue (i.e. structurally similar) of the target peptide (Sternberg *et al.* page 371, column 1, lines 10-19). Thus, in some instances, the *ab initio* methods were able to predict structures better than methods that used knowledge of polypeptide similarity. Contrary to the Examiner's assertion, Sternberg *et al.* does not disclose that *ab initio* methods have a poor ability to predict protein structure.

Table 1 in Koehl *et al.* shows the best results of the structure prediction of CASP3 targets based on comparative modeling, fold recognition, and *ab initio* methods. Although some *ab initio* methods predicted the positions of only about half of the residues in the sequence, Koehl *et al.* does not disclose that *ab initio* algorithms per se have a poor ability to predict protein structure. In fact, Koehl *et al.* discloses that the quality of models for the five easiest *ab initio* targets were as good as that of the four most difficult folding recognition targets (Koehl *et al.* Table 1 and page 109, column 3, lines 52-56). For example, the prediction of target 56 as a new fold was close to the native structure of the

target peptide with a Cα RMS of 6Å (Table 1). **Additionally, Koehl *et al.* also discloses that, although the CASP assesses the ability of the best practitioners in the field to predict protein structure, it does not test how well other scientists can expect to do, or how well totally automated methods will do in *ab initio* structure prediction (page 111, column 2, 5-10).** Therefore, Koehl *et al.* does not disclose that *ab initio* methods have a poor ability to predict protein structure and does not disclose that *ab initio* methods cannot be used to determine 3-D structures of proteins.

Further, as noted above, *ab initio* methods that can be used to generate 3-D protein structure were available at the effective time of filing of the instant application. For example, as disclosed in the specification, the Balaji *et al.* methods can be used to generate 3-D protein structures without any regard to the conformation of the underlying primary or secondary structure. As noted above, the Balaji *et al.* methods generated 3-D polypeptides models of endothelin-1 that were good enough to design peptide analogs therefrom that mimic the surface features endothelin-1 and modulate the activity of endothelin-1. Since the Balaji *et al.* methods generated polypeptide models that were good enough to design bioactive analogs, the Balaji *et al* methods are good enough to design polypeptide models that can be used in the computer-based methods of drug design described in the instant application. The Examiner has not set forth any reasons or evidence demonstrating that the *ab initio* methods of Balaji et al. are not suitable for the instantly claimed methods.

2. Rebuttal to the Examiner's predictability arguments

The Examiner alleges that Sternberg *et al.* and Koehl *et al.* show that *ab initio* methods of structure prediction from polypeptide sequence information alone is not predicted to result in an accurate structure of a complete polypeptide.

As noted above, Sternberg *et al.* discloses that *ab initio* methods generated good 3-D protein models from protein sequences alone without employing fold recognition and that, in some instances, *ab initio* methods were able to predict structures better than methods that used knowledge of polypeptide similarity. (Sternberg *et al.* page 371, column 1, lines 10-19, and column 2, lines 31-34). Sternberg *et al.* does not disclose that *ab initio* methods in general have a poor ability to predict accurate structures of complete polypeptides.

Despite the fact that some *ab initio* methods predicted the positions of only about half of the residues in the sequence in Koehl *et al.*, Koehl *et al.* does not disclose that *ab initio* algorithms in general have a poor ability to predict accurately structures of complete polypeptides. **In fact, Koehl *et al.* discloses that the CASP does not test how well other scientists can expect to do or how well totally automated methods will do in *ab initio* structure prediction (page 111, column 2, 5-10).** Since the reference can not be used to predict how well other scientists will do using *ab initio* methods, the reference certainly can not be construed to mean that all *ab initio* methods will result in inaccurate predictions of the structures of complete polypeptides. In fact, as noted above, the Balaji *et al. ab initio* methods have been used to generate structures of complete polypeptides that are accurate enough to design drug analogs that are bioactive. The Balaji *et al. ab initio* methods described above can be used to generate structures of complete polypeptides that are accurate enough for use in the drug design methods of the instant application.

Assuming arguendo that *ab initio* methods are inaccurate in predicting structures of complete polypeptides, other methods can be used alone or in combination with *ab initio* methods to generate 3-D structures of target proteins as discussed above.

Assuming arguendo that *ab initio* methods are inaccurate in predicting structures of complete polypeptides *ab initio* methods can accurately predict the structures of portions of polypeptides, as disclosed in Koehl *et al.* and noted by the Examiner, and these structures can be used in the instantly claimed methods. For example, in the methods of computer-based drug design claim in the instant application, 3-D structures of target proteins that are the product of genes exhibit genetic polymorphisms are determined and used to design, modify, or identify candidate drugs that interact with the 3-D structures. The 3-D structures that are determined can be whole structures of the target proteins or partial structures of the target protein. As taught in the specification, the drug design methods of the instant application can be used to design drugs that interact with structural variants **or portions thereof** by generating 3-D protein structural variant models of **all or a portion** of the protein (page 21, lines 15-30). Thus, portions of protein structures can be used in drug design. Therefore, if the portions of proteins that are generated accurately by *ab initio* methods are active sites or sites that are targeted by drugs, then even those portions of proteins can be used in the drug design methods of the instant application.

Assuming arguendo that *ab initio* methods have a poor ability to predict protein structure without structural information from other structurally similar polypeptides, *ab initio* methods can be used to generate 3-D protein structures that can be used in the instant methods of computer-based drug design discussed above. For example, protein structures generated solely by *ab initio* methods can be validated by comparing data from computational binding studies to data from experimental binding studies, as discussed above. Models that have been validated by this comparison can be used in the instant methods of computer-based drug design as shown in the Examples.

Therefore, the accuracy of *ab initio* methods in generating 3-D protein structures is largely irrelevant because *ab initio* methods can be used to generate portions of 3-D protein structures accurately and can be used to

PATENT & TRADEMARK
JUN 0 3 2003
OIPE

generate 3-D protein structures that can be validated by experimental binding studies. The 3-D protein structures or portions thereof generated by the *ab initio* methods can be used in the instantly claimed methods.

Furthermore, *ab initio* methods, such as the Balaji *et al. ab initio* methods, can be used to generate 3-D protein structure that can be used in the instantly claimed methods, as noted above. These *ab initio* methods were available at the effective time of filing of the instant application.

* * *

Respectfully submitted,
HELLER EHRMAN WHITE & McAULIFFE LLP

By: _____
Stephanie L. Seidman
Registration. No. 33,779

Attorney Docket No. 24737-1906C
Address all correspondence to:
Stephanie L. Seidman, Esq.
HELLER EHRMAN WHITE & McAULIFFE
4350 La Jolla Village Drive, 7th Floor
San Diego, California 92122-1246
Telephone: 858 450-8400
Facsimile: 858 587-5360
email:sseidman@HEWM.com

PATENT & TRADEMARK OFFICE
JUN 03 2003
OIPE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Ramnarayan *et al.*
Serial No.: 09/709,905
Confirmation No.: 3606; cust. no. 24961

Filed: November 10, 2000

For: *USE OF COMPUTATIONALLY DERIVED PROTEIN STRUCTURES OF GENETIC POLYMORPHISMS IN PHARMACOGENOMICS FOR DRUG DESIGN AND CLINICAL APPLICATIONS*

Art Unit: 1631

Examiner: Brusca, J.

ATTACHMENTS TO THE AMENDMENT

1. Marked-up specification paragraphs and claims
2. Marked-up Figures 4, 6 and 7
3. Copies of formal drawings of Figures 4, 6, 7 and 11
4. The following documents:

Dudek *et al.* "Protein Prediction Using a Combination of Sequence Homology and Global Energy Minimization: II. Energy Functions," (1998) J. Comput. Chemistry 19:548-573;

Abagyan and Totrov, "Ab Initio Folding of Peptides by the Optimal-Bias Monte Carlo Minimization Procedure," (May 1, 1999) 151:402-421;

Osguthorpe, "Improved Ab Initio Predictions with a Simplified, Flexible Geometry Model," Proteins: Structure, Function, and Genetics Suppl 3 (November 9, 1999) 186-193;

Westhead and Thornton "Protein structure prediction," Curr Opin in Biotechnology (1998) 9:383-389;

Eisenhaber *et al.* "Protein structure prediction: recognition of primary, secondary, and tertiary structural features from amino acid sequence," Critical Rev. in Biochem and Mol. Biol. (1995) 30:1-94;

Jones, "Successful *ab initio* prediction of the tertiary structure of NK-Lysin using multiple sequences and recognized supersecondary structural motifs," Proteins: Structure, function, and Genetics, Suppl 1 (1997) 185-191;

Samudrala *et al.*, "Ab initio protein structure prediction using a combined hierarchical approach," Proteins: Structure, function, and Genetics Suppl 3 (1999) 194-198;

Dunbrack *et al.* "Meeting review: the Second Meeting on the Critical Assessment of Techniques for Protein Structure Prediction (CASP2), Asilomar, California, December 13-16, 1996," Folding and Design (1997) R27-R42; and

de Dios *et al.* "Secondary and Tertiary Structural Effects on Protein NMR Chemical Shifts: An ab Initio Approach," Science (1993) 260:1491-1496.

U.S. Patent No. 5,736,509 (April 7, 1998)

HAND-ANNOTATED COPY OF FIGURES 4, 6, AND 7